

Europäisches Patentamt  
European Patent Office  
Office européen des brevets



(11)

EP 0 974 933 A2

(12)

## EUROPEAN PATENT APPLICATION

(43) Date of publication:  
26.01.2000 Bulletin 2000/04

(51) Int Cl.7: G06T 9/00

(21) Application number: 99304722.4

(22) Date of filing: 16.06.1999

(84) Designated Contracting States:  
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU  
MC NL PT SE  
Designated Extension States:  
AL LT LV MK RO SI

(72) Inventors:  
• Konstantinides, Konstantinos  
San Jose, California 95118 (US)  
• Tretter, Daniel R.  
No.18, Mountain View, California 94040 (US)

(30) Priority: 21.07.1998 US 120072

(74) Representative: Powell, Stephen David et al  
WILLIAMS, POWELL & ASSOCIATES  
4 St Paul's Churchyard  
London EC4M 8AY (GB)

(71) Applicant: Hewlett-Packard Company  
Palo Alto, California 94304 (US)

### (54) ADAPTIVE VIDEO COMPRESSION USING VARIABLE QUANTIZATION

(57) An image compression system [50] for compound images containing both text and pictures is capable of receiving the images on a non-overlapping 8 by 8 pixel block and includes a discrete cosine transformer [14] connected to a quantizer [18] drawing lossy quantization factors from quantization tables [16]. The lossy quantization factors are modified by a variable quantization subsystem [54] based on the frequency of

changes in the block to provide low lossy quantization factors for high frequency of changes and high lossy quantization factors for low frequency of changes, the high frequency of changes being indicative of text and the low frequency of changes being indicative of pictures. The quantizer [18] is connected to an entropy coder using lossless entropy encoding factors from Huffman tables [22] to provide JPEG compliant files [26].

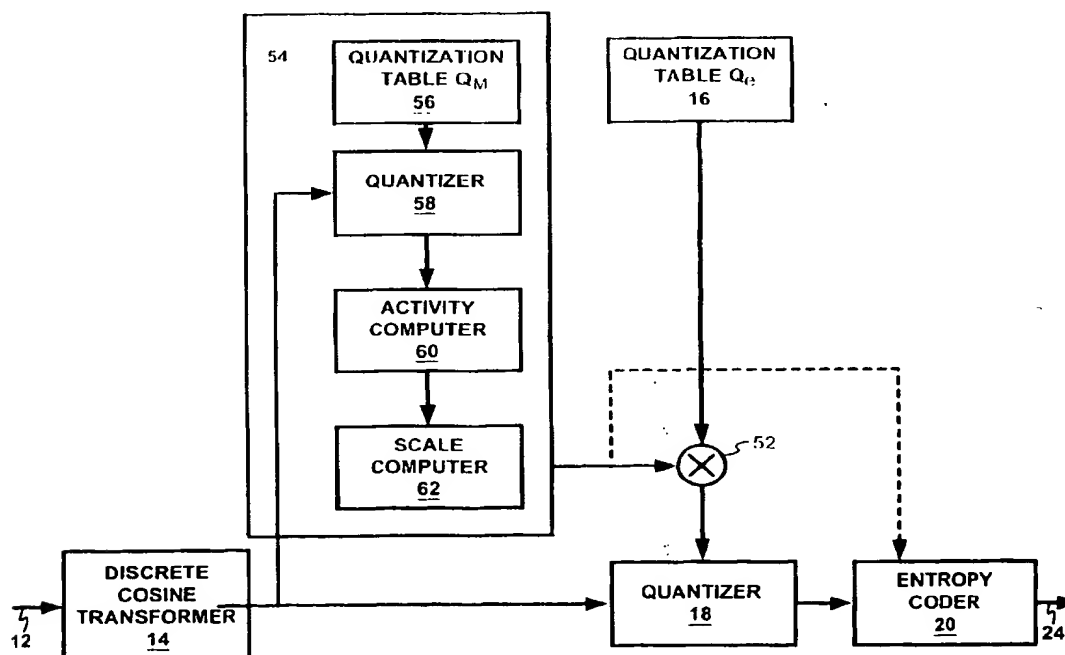


FIG. 3

EP 0 974 933 A2

## Description

[0001] The present invention relates generally to a system for variable quantization in JPEG for compound documents and more specifically to JPEG data compression for compound images having pictures and text.

[0002] JPEG is the name of both a committee and a standard JPEG stands for joint Photographic Experts Group, the original name of the committee that wrote the JPEG standard. The JPEG standard is an international standard which applies to the lossy and lossless compression of either full-color or gray-scale images of natural, real-world scenes.

[0003] Lossy image compression compresses by striving to discard as much of the image data as possible without significantly affecting the appearance of the image to the human eye. Lossless compression is compression achieved without discarding any of the image data.

[0004] The JPEG standard works well on still photographs, naturalistic artwork, and similar material (which are generally referred to herein as "pictures"), but not so well on lettering simple cartoons, or line drawings (which are generally referred to herein as "text"). Compound images are those which contain both pictures and text (which are collectively referred to herein as "images"). In some cases, compound images contain pictures which also contain text within the picture itself.

[0005] This standard is being used in the computer industry. Popular graphics-capable browsers on the World Wide Web can read and write this particular type of image data format. so if a compressed image is sent across the Web to such a browser, it knows how to decompress the image and display it.

[0006] Compression is important for two main reasons. The first is storage space. If there will be a large number of images on a hard drive, the hard drive will fill up very quickly unless the data can be greatly compressed. Computers have fixed size buffers and limited memory, and an image has to fit in them otherwise, the image cannot be stored in them.

[0007] The second is bandwidth. If data is being sent through a browser or through electronic mail, the more bits that need to be transmitted, the more time is required. For example, with a 28.8K modem it may take half an hour of waiting for a picture to be completely transmitted. If a 50 to 1 compression can be achieved, the same picture can be transmitted completely in about thirty seconds, and if compressed properly, the recipient will not notice the difference between the original and the compressed version.

[0008] For full-color images, the uncompressed data is normally 24 bits per pixel, JPEG can typically achieve 10:1 to 20:1 compression on pictures without visible loss, bringing the effective storage requirement down to 1 to 2 bits per pixel. This is due to the fact that small color changes are perceived less accurately than small changes in brightness. Even 30:1 to 50:1 compression is possible with small to moderate defects, while for very low quality purposes such as previews or archive indexes, 100:1 compression is quite feasible.

[0009] For gray-scale, and black and white images such large factors of compression are difficult to obtain because the brightness variations in these images are more apparent than the hue variations. A gray-scale JPEG file is generally only about 10%-25% smaller than a full-color JPEG file of similar visual quality with the uncompressed gray-scale data at only 8 bits/pixel, or one-third the size of the color data. The threshold of visible loss is often around 5:1 compression for gray-scale images.

[0010] Although there are a number of settings that can be predefined to achieve different compression ratios, there is only one parameter, called the quality factor, that is adjusted regularly in JPEG on an image-by-image basis with one setting for an active image. The quality factor is a single number in an arbitrary, relative scale. A high quality factor will provide a relatively high quality decompressed image, but will require a relatively large file. And, of course the lower the quality, the rougher the approximation of the image and the more compression with a correspondingly smaller file size, but also, the more visible defects, or artifacts, will be in the decompressed final image. Text generally shows significant compression artifacts at higher quality factors than pictures. Further, the quality factor will only give an approximate end file size.

[0011] Therefore, a long sought goal in image compression has been to maintain maximum perceptible image quality while achieving maximum compression.

[0012] This goal is becoming more difficult to attain because compound documents are just starting to become more and more important. It has only been recently that it has become possible to drop pictures into text documents as much as can be done now. Before, electronic transmissions were either a text document or a picture document. Now, it is more and more common to see a compound image where someone is making a newsletter or setting up a website. People want to drop in some pictures but also want to have text as well. So compound documents are becoming a more important, whether it is just photocopying or just sending to a printer or transmitting across the internet, these have become a more important class of images.

[0013] Also, most of the techniques that have been developed in the past for compound documents are based on proprietary (non-standard) compression techniques, so the images could only be decompressed using a specific company's product.

[0014] It has long been known that the inability to minimize file size while maintaining high perceptual quality would lead to detrimental compromises in performance so process improvements have been long sought but have eluded those skilled in the art. Similarly, it has long been known that the problems would become more severe with compound documents and thus a generally applicable solution has been long sought.

[0015] The present invention provides a simple metric for picture, text segmentation of compound documents in the discrete cosine transform domain. This allows areas of high frequency content such as text to be compressed at a better quality than pictures, thus improving the overall perceptual quality while minimizing the file size. The metric is computed using the quantized output of the discrete cosine transform. No other information is needed from any other part of the JPEG coder.

[0016] The present invention provides an image compression system which can be used to apply different, appropriate quantization factors to small blocks of pictures and text to provide significant image compression.

[0017] The present invention further provides an image compression system capable of distinguishing between text and pictures in compound images.

[0018] The present invention still further provides for preserving the text quality without sacrificing bandwidth while at the same time being JPEG compliant.

[0019] The present invention also provides an image compression system which is fully compliant with the latest extensions of the current JPEG standard.

[0020] The above and additional advantages of the present invention will become apparent to those skilled in the art from a reading of the following detailed description when taken in conjunction with the accompanying drawings.

FIG. 1 is a schematic view of a prior art baseline JPEG encoder;

FIG. 2 is a schematic view of JPEG Part 3 encoder that supports variable quantization; and

FIG. 3 is a schematic view of a variable quantization subsystem of the present invention.

[0021] Referring now to FIG. 1 PRIOR ART, therein is shown a baseline JPEG encoder system 10 for digital cameras, scanners, printers, imaging servers, etc. The JPEG encoder system 10 is for an image with a single color component. For color images, there would be a JPEG encoder system 10 for each of the color components.

[0022] The system 10 receives image pixels, or input digital image data, at an input 12 which is connected to a discrete cosine transformer 14. The discrete cosine transformer 14 first divides the input digital image data 12 into non-overlapping, fixed length image blocks, generally 8 by 8. After a normalization step, the discrete cosine transformer 14 reduces data redundancy and transforms each fixed length image block by applying a discrete cosine transform to a corresponding block of discrete cosine transform coefficients. This transform converts each fixed length image block into the frequency domain as a new frequency domain image block. The first coefficient in the block, the lowest frequency coefficient, is the DC coefficient and the other coefficients are the AC coefficients (e.g., for an 8 by 8 block, there will be one DC coefficient and 63 AC coefficients).

[0023] Quantization tables 16 are operatively connected to the discrete cosine transformer 14. The quantization tables 16 contain lossy quantization factors (scaled according to the factor) to be applied to each block of discrete cosine transform coefficients. One set of sample tables is given in Annex K of the JPEG standard (ISO/IEC JTC1 CD 10918:ISO, 1993). These tables and the user-defined quality factors do not actually provide compression ratios per se, but provide factors indicating how much the image quality can be reduced on a given frequency domain coefficient before the image deterioration is perceptible.

[0024] It should be understood that the tables represent tabulations of various equations. The look-up tables could be replaced by subroutines which could perform the calculations to provide the factors.

[0025] A quantizer 18 is connected to the discrete cosine transformer 14 and the quantization tables 16 to divide each frequency domain image block by the corresponding element from the quantization table 16 to output the quantized discrete cosine transform output.

[0026] An entropy coder 20 is connected to the quantizer 18 and to Huffman tables 22. The entropy coder 20 receives the output from the quantizer 18 and rearranges it in zigzag order. The zigzag output is then compressed using run-length encoding in the entropy encoder 20 which is a lossless entropy coding of each block of quantized discrete cosine transform output. The entropy encoder 20 of the present invention uses Huffman codes from the Huffman tables 22 although arithmetic coding can also be used. The Huffman codes exploit similarities across the quantized discrete cosine transform coefficients. JPEG contains two sets of typical Huffman tables, one for the luminance or grayscale components and one for the chrominance or color components. Each set has two separate tables, one for the DC components and the other for the AC.

[0027] The bitstream out of the entropy coder 20 at output 24 is a JPEG file 26 which contains headers 28, tables 30, and data 32. The tables 30 contain information from the quantization tables 16 and the Huffman tables 22 of the appropriate information used in the processing of each block of data so the data can be properly decompressed. The data 32 contains the output from the entropy coder 20 in a form of a compressed block such that a sequence of all of

the compressed blocks forms the compressed digital image data.

[0028] Referring now to FIG. 2, therein is shown a JPEG encoder system 50 that supports the variable quantization of the present invention. The system 50 is compliant with the JPEG Part 3 standard. The same elements as in FIG. 1 are given the same numbers in FIG. 2. Thus, the system 50 receives input digital image data 12 into the discrete cosine transformer 14 which is connected to the quantizer 18.

[0029] The quantization tables 16 are connected to a multiplying junction 52 which is connected to the quantizer 18. Also connected to the multiplying junction 52 is a variable quantization subsystem 54 which is also connected to the entropy coder 20.

[0030] The entropy coder 20 is connected to the quantizer 18 and to the Huffman tables 22. The bitstream out of the entropy coder 20 at output 24 is a JPEG file 26 that contains headers 28, tables 30, and data 32. The tables 30 contain compression-related information from the quantization tables 16 and the Huffman tables 22 which can be used by a JPEG decompressor system (not shown) to decompress the data 32 from the entropy coder 20. The quantization scale factor from the variable quantization subsystem 54 are incorporated into the data 32 by the entropy encoder 20.

[0031] Referring now to FIG. 3, therein is shown the variable quantization subsystem 54 which is operatively connected to the discrete cosine transformer 14. The discrete cosine transformer 14 is connected to a quantizer 58 in the variable quantization subsystem 54. The quantizer 58 has quantization tables 56 connected to it which are the factors relating to the activity metrics. The quantizer 58 is the same as the quantizer 18. For simplicity, the quantization tables 56 for the activity metrics are the same as the quantization tables 16 for the encoding, but this is not necessary.

[0032] The quantizer 58 is further connected to an activity computer 60 for computing an activity metric,  $M_i$ , as will later be described. The activity computer 60 is connected to a scale computer 62 for computing  $qscale$  as will also later be described. The scale computer 62 is connected to the multiplying junction 52 to which the quantization table 16 is connected. The discrete cosine transformer 14 as well as the multiplying junction 52 are connected to the quantizer 18 as also shown in FIG. 2.

[0033] In operation in the FIG. 1 PRIOR ART baseline JPEG encoder system 10, image pixels are divided into non-overlapping  $8 \times 8$  blocks where  $y_i$  denotes the  $i$ -th input block. This division applies to all images including text, pictures, and compound images as well as pictures containing text. After a normalization step, each block is transformed into the frequency domain using the discrete cosine transform.

[0034] Mathematically JPEG uses the discrete cosine transform in its processing. A discrete cosine transform in the frequency domain is based on the assumption that an image mirrors itself on both boundaries. This assures a smooth transition without high frequency spikes, because those are very hard to compress. And the higher frequencies are very close to zero if there is a smoothly varying function.

[0035] The output of the above step will be a new  $8 \times 8$  matrix  $Y_i$ . Next, each element of  $Y_i$  is divided by the corresponding element of the encoding quantization table  $Q_e$ . Given  $Q_e[j, k]$ ,

$$Y_{Q_e}[j, k] = \text{round} \left( \frac{Y_i[j, k]}{Q_e[j, k]} \right)$$

[0036] In the baseline JPEG encoder system 10, a single set of quantization tables 16 is used for the whole image. For a *compound* 1 ISO test image, the whole image would be 512 pixels by 513 pixels.

[0037] After quantization, the quantized discrete cosine transformer 14 output is rearranged in raster, or zigzag, order and is compressed using run-length encoding which is a lossless, entropy coding using the Huffman tables 22 in the entropy coder 20. According to the JPEG standard, the quantization tables for each color component can be defined in the header tables 30 of the JPEG file 26. The output of the entropy coder 20 is a sequence of compressed blocks which form the compressed image which can be decompressed by a standard JPEG decompression system.

[0038] In operation, the variable quantization JPEG encoder system 50 of the present invention shown in FIG. 2 uses the JPEG adjunct standard (ISO/ IEC JTC1 CD 10918; ISO, 1993) which has been extended to support variable quantization. Under variable quantization, the values of the original quantization matrix can be rescaled on small blocks of pixels as small as 8 pixels by 8 pixels. Normally the quantization matrix stays the same for the entire image, but the adjunct standard allows these changes on a block by block basis. This change was designed primarily for a rate control problem of getting the proper number of bits at the output. If a block is changed, the information can be put into the bitstream so that a decoder on the receiving end can undo it later. Thus, the various scaling factors are also encoded as part of the data bitstream. In principal, variable quantization allows for better rate control or more efficient coding which is the original purpose for the JPEG extension.

[0039] Even though the latest JPEG extensions provide the syntax for the support of variable quantization, the actual

way to specify the scaling factors is application dependent and not part of the JPEG adjunct standard.

**[0040]** The present invention allows for the variable quantization of compound images. The JPEG encoder system 50 of FIG. 2 automatically detects the text-part and the image-part of a document by measuring how quickly pixels are changing in the incoming data. Black text on a white background changes very quickly from black to white and back again even within a very small block of pixels. Picture pixels, for example the image of a human face, change much more slowly through gradations of color.

**[0041]** In the JPEG standard, the quantization is being done in a transformed domain, not on the pixels directly. The pixels are transformed through a linear matrix transform (the discrete cosine transform) into a frequency domain representation, and the quantization is performed in the frequency domain. It is also in the frequency domain that the frequency components can be determined to find how active a particular block is. The mathematical equation used is what provides an "activity metric". The larger this activity metric turns out to be, the more things are changing within the 8 by 8 block.

**[0042]** Also, the discrete cosine transform has the advantage that it takes real numbers and transforms them into real numbers which can be quantized. In this domain, it is possible to predict, roughly, how many bits it takes to represent the data that is actually in the block.

**[0043]** To represent a given number, such as every number between 0 and 15, the largest number is taken, and the log base 2 of this provides the number of bits required. In this case it would be 4 because with 4 bits every number between 1 and 15 can be represented.

**[0044]** Thus by taking the absolute value of the real numbers, taking the log base 2 of them, which tells how many bits needed for each one and then summing them up, and that provides how many bits are needed to represent the data in the entire block. This number is going to be very large if there is a lot of activity because a lot of the frequency components will be big, and it is going to be very small if the image changes slowly because all the high frequencies will be close to zero.

**[0045]** Based on the discrete cosine transform activity of a block or a macroblock (a 16 x 16 block), quantization scaling factors are derived that automatically adjust the quantization so that text blocks are compressed at higher quality than image blocks. Those skilled in the art would be aware that text is more sensitive to JPEG compression because of its sharp edges which, if compressed too greatly, would blur or have ringing artifacts (ripples around the edges). At the same time, images can be compressed greatly without drastically affecting human-eye perceived differences in quality of the image.

**[0046]** There are many ways to measure activity in a block. One method to determine the discrete cosine transform activity is to let  $Y_i[j, k]$  denote the elements of the  $i$ -th output block of the discrete cosine transform so  $Y_i[0, 0]$  denotes the DC component of the  $i$ -th block. In the present invention, the activity computer 60 uses the following activity metric:

$$M_i = \frac{1}{64} \left[ \log_2 Y_i[0,0] - Y_i[0,0] + \sum_{j,k} \log_2 Y_i[j,k] \right]$$

where the summation is performed over all elements of the  $Y_i[j, k]$  matrix, except  $Y_i[0, 0]$ . (The above formulation assumes that arguments in the  $\log_2$  function are always greater than zero.)

**[0047]** The motivation behind this metric is based on the fact that: (a) JPEG uses differential coding for the encoding of the discrete cosine coefficients, and (b) the number of bits needed to code a DC transform coefficient are proportional to the base-two logarithm of its magnitude. The equation does not (and does not need to) account for additional coding bits needed for the Huffman coding of either size or run/size information of the discrete cosine transform coefficients. The number of these bits ranges from 2 to 16 per non-zero coefficient. Assuming that on the average  $c$  bits per non-zero coefficient are required for this purpose, then the following method can be used to compute  $M_i$  (and the overall bit rate) more accurately.

Begin

$M_i = 0$

$D = |Y_i[0, 0] - Y_{i-1}[0, 0]|$

if  $D > 0$  then

$M_i = M_i + \log_2 D + c$

$j = 0$

for  $k = 1$  to 7

if  $|Y_i[j, k]| > 0$  then

$M_i = M_i + \log_2 |Y_i[j, k]| - c$

for  $j = 1$  to 7

for  $k = 0$  to 7

if  $|Y_i[j, k]| > 0$  then

$M_i = M_i + \log_2 |Y_i[j, k]| + c$

End

In experiments,  $c = 4$ .

**[0048]** After defining an activity metric, the next step is to define the relationship between the activity measure and the quantization scale.

**[0049]** The *compound* ISO test image is used as the standard compound image with computer generated text, a photographic-type color image, and computer generated text within the image. The top half is text and the bottom half the color image. This is a 512 x 513 pixels total image, with 1,056 macroblocks (16 x 16 pixel blocks). The color image part starts at approximately the 508-th macroblock.

**[0050]** The values of the activity metric,  $M_i$ , calculated in the activity computer 60 for each of the luminance macroblocks in the ISO test image is higher in the text regions of the image. However, discrimination between image and text areas is even better if  $M_i$  is computed using quantized values of  $Y_i$ ; that is,  $Y_{QM,i}$ . When the quantization matrices  $Q_M$  and  $Q_C$  in quantization tables 56 and 16, respectively, are both the same as the one given in Annex K of the JPEG standard, activity values larger than 1.2 correspond to text areas in the image. It should be understood that the two quantization tables 56 and 16 could be different. Experiments show that the range of values for  $M_i$  for the ISO test image is consistent with the range of values obtained from other test images.

**[0051]** Basically quantization varies inversely with the metric. If a higher metric means a higher activity, thus scale by less to quantize more finely or compress less. And then with a very small metric, scale the quantization very coarsely or compress more because the image is such a smooth block, that it does not matter how much quantization since it will not be perceptible.

**[0052]** The output of the scale computer 62 is *qscale*, which denotes the parameter used to scale the AC values of the original quantization matrix, a value of *qscale* = 0.5 is quite acceptable for the compression of text. On the other hand, values of *qscale* larger than 2 may yield serious blocky artifacts on an image. To simplify implementation, a linear, but bounded, relationship between *qscale* and the activity metric ( $M_i$ ) is superior, such as

$$qscale_i = \begin{cases} a \times M_i + b & \text{if } 2 > a \times M_i + b \geq 0.4 \\ 0.4 & \text{if } a \times M_i + b < 0.4 \\ 2 & \text{if } a \times M_i + b > 2 \end{cases}$$

where  $a$  and  $b$  are constants to be defined based on desired output quality and compression ratios. One way  $a$  and  $b$  can be defined is the follows. Let  $m_l$  denote the value of the activity metric,  $M_l$ , for which  $qscale = 1$ . Let  $m_u$  denote the value of  $M_l$  for which  $qscale = 0.5$ . After solving two equations with two unknowns:

$$a = \frac{0.5}{m_l - m_u}$$

$$b = 1 - m_l \times a$$

For example, if  $m_l = 0.6$  and  $m_u = 1.2$ , then  $a = -0.83$  and  $b = 1.498$ .

**[0053]** The choices for  $m_l$  and  $m_u$  effect compression ratios as follows. If  $m_l$  is increased, in effect more blocks are quantized with a  $qscale > 1$ ; thus compression is improved but image quality may be reduced. If  $m_u$  is increased, the number of blocks that are quantized with  $qscale = 0.5$  is decreased; thus the quality of text is decreased, but the compression ratios are improved.

**[0054]** In the variable quantization subsystem of FIG. 3 showing the variable quantization method, the  $Q_M$  quantization matrix is the same as  $Q_e$ , but this may not be always the case. For example,  $Q_M$  may be the same as Appendix K of the JPEG standard, but  $Q_e$  may be a custom quantization table.

**[0055]** Using the above metric, the upper text areas in the ISO test image were identified as areas of high frequency activity, but also the text inside the color picture at the bottom half.

**[0056]** The  $qscale$  is provided to the multiplying junction 52 to control  $Q_e$  from the quantization table 16 to the quantizer 18.

**[0057]** It should be understood that the same method as described above may be used to adjust the chroma quantization tables independently from the luminance tables using the same scaling factors.

**[0058]** JPEG, itself, as a standard doesn't specify what to do about color. But what is commonly done, is to convert a color image into a luminance and chrominance representation so that it shows the brightness of the image with two other components that show the colorfulness. And it turns out that the human eye is much more sensitive to the luminance. A slow transition from a red to an orange versus a sharp one will not even be noticeable, but a slow transition in luminosity versus a sharp one will be noticeable as a blur. In the present invention, the activity metric is computed only for the luminance component to save on computation and then the chrominance is scaled the same way. That turns out to work reasonably well on the chrominance as well, because compound documents usually have black text on a white background and errors in the color are particularly noticeable. A little red fringe around each letter will be seen immediately but it is less likely to be seen in an image.

**[0059]** As would be understood by those skilled in the art, the present invention has been described in terms of discrete components but it may be carried out in software or in dedicated integrated circuits.

**[0060]** While the invention has been described in conjunction with a specific best mode, it is to be understood that many alternatives, modifications, and variations will be apparent to those skilled in the art in light of the foregoing description.

## Claims

1. Variable quantization apparatus [50] for an image encoder system having an interconnected transformer [14], quantizer [18], and entropy encoder [20], comprising:

variable quantization means [54] operatively connected to the transformer [14] and the quantizer [18] responsive to plurality of blocks of data from the transformer [14] to determine the characteristics of a plurality of blocks of digital pixel data inputted to the transformer [14] by computing an image related metric for each of

the plurality of blocks of data;  
 quantization factoring means [16] connected to said variable quantization means [54] for providing a predetermined lossy quantization factor for each metric; and  
 said variable quantization means [54] including means [58] for causing the quantizer [18] to apply a lossy quantization factor to each of the plurality of blocks of data based on said metric for the block of data to provide a plurality of blocks of quantized data to the entropy encoder [20].

2. The variable quantization apparatus [50] as claimed in claim 1 wherein:

said variable quantization means [54] includes scaling means [62] for scaling said metric for each of the plurality of blocks of data to provide a lower lossy quantization factor for metrics for predetermined types of images and a higher lossy quantization factor for metrics for other predetermined types of images.

3. The variable quantization apparatus [50] as claimed in claim 1 wherein: said variable quantization means [54] includes:

quantization tables [56] for providing a predetermined lossy quantization factor for predetermined of blocks of data;  
 a quantizer [58] connected to the transformer [14] and said quantization tables [56] for applying a lossy quantization factor to said each of said plurality of blocks of data to compute said metric for each of said plurality of blocks of data.

4. The variable quantization apparatus [50] as claimed in claim 1 including:

multiplier means [52] connected to said quantization tables [16], said variable quantization means [54], and said quantizer [18] for applying a lossy quantization factor which is a function of said metric and data to said each of said plurality of blocks of data.

5. The variable quantization apparatus [50] as claimed in claim 1 wherein:

said variable quantization means [54] computes the metric as an image activity metric according to the equation:

$$M_i = \frac{1}{64} \left[ \log_2 |Y_i[0,0] - Y_{i-1}[0,0]| + \sum_{j,k} \log_2 |Y_i[j,k]| \right]$$

where:

$Y_i[j,k]$  denote the elements of the  $i$ -th block output by the transformer; and  
 $Y_i[0,0]$  denotes the transformed DC component of the  $i$ -th block.

6. The variable quantization apparatus [50] as claimed in claim 1 wherein:

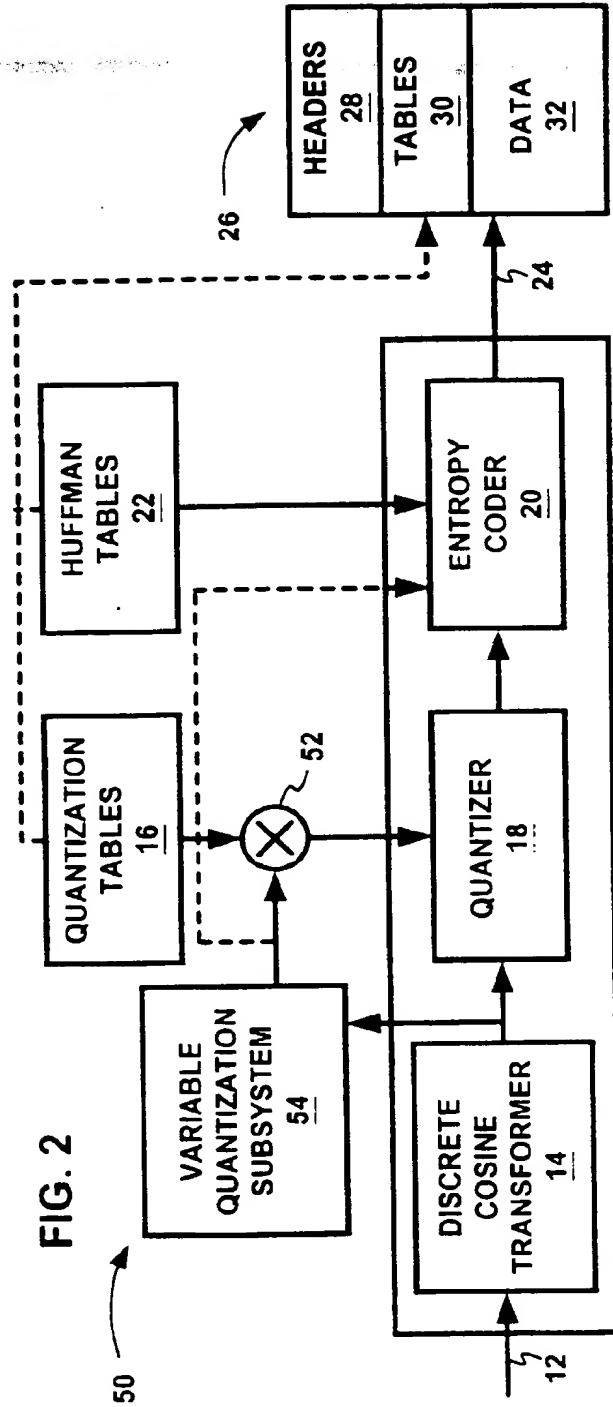
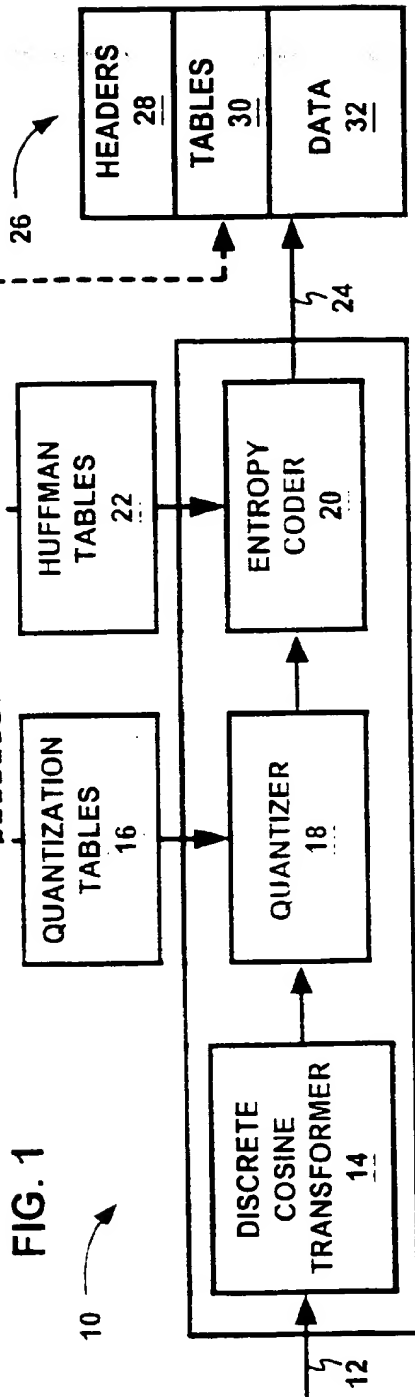
said variable quantization means [54] includes scaling means [62] for scaling said metric as an activity matrix for each of the plurality of blocks of data provided by the transformer [14] as frequency data to provide a lower lossy quantization factor for higher activity metrics proportionally than a higher lossy quantization factor for lower activity metrics whereby one type of image is compressed at a higher quality while another type of image is compressed at a lower quality according to the equations:



$$qscale_i = \begin{cases} a \times M_i + b & \text{if } 2 > a \times M_i - b \geq 0.4 \\ 0.4 & \text{if } a \times M_i - b < 0.4 \\ 2 & \text{if } a \times M_i - b > 2 \end{cases}$$

where:  $a$  and  $b$  are predetermined functions of said activity metric.

7. The variable quantization apparatus [50] as claimed in claim 1 wherein said blocks of digital pixel data are 8 by 8 pixels.



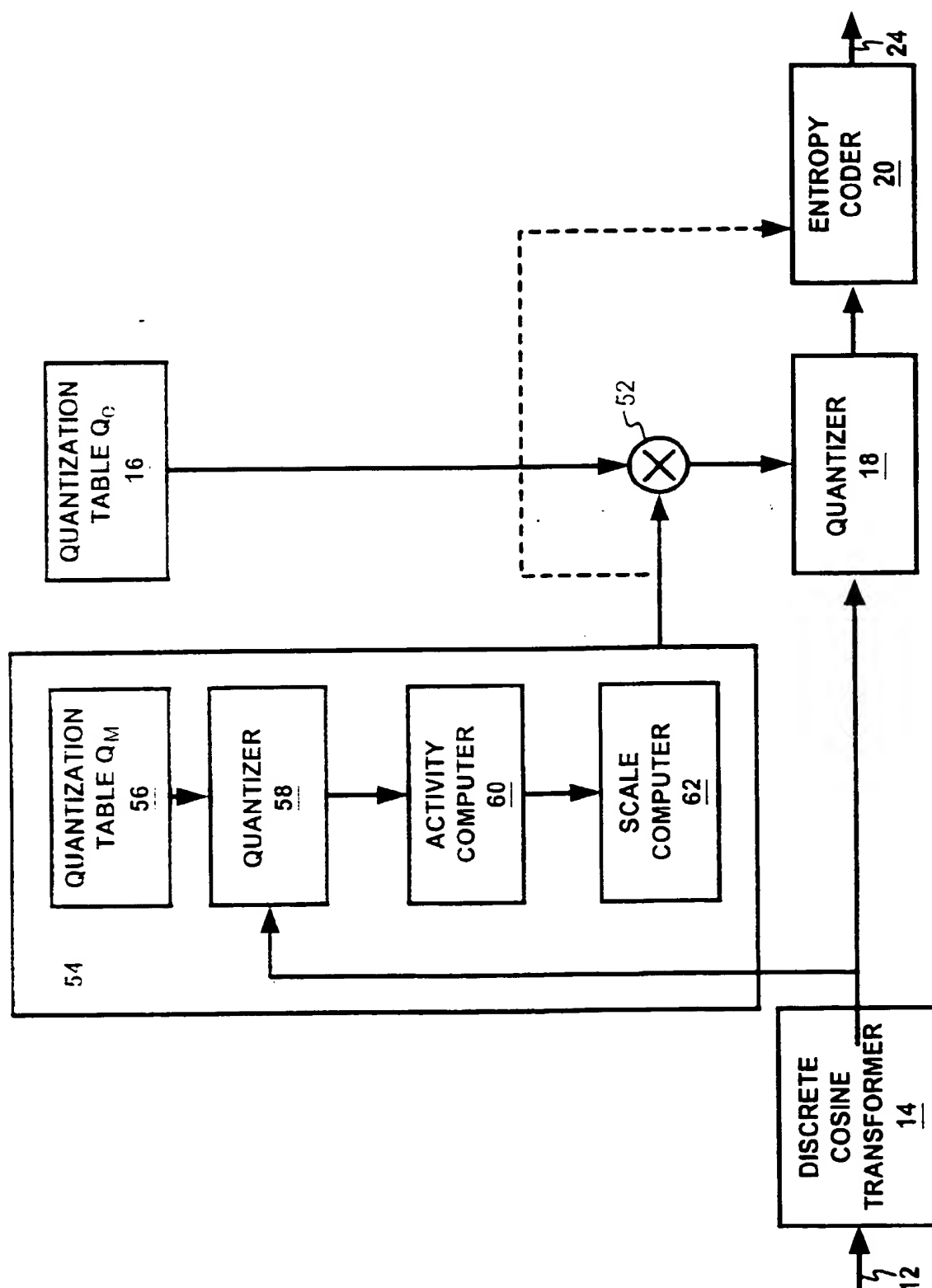
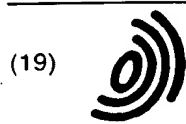


FIG. 3

**THIS PAGE BLANK (USPTO)**



Europäisches Patentamt  
European Patent Office  
Office européen des brevets



(11)

EP 0 974 933 A3

(12)

## EUROPEAN PATENT APPLICATION

(88) Date of publication A3:  
03.01.2001 Bulletin 2001/01

(51) Int Cl.7: G06T 9/00

(43) Date of publication A2:  
26.01.2000 Bulletin 2000/04

(21) Application number: 99304722.4

(22) Date of filing: 16.06.1999

(84) Designated Contracting States:  
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU  
MC NL PT SE  
Designated Extension States:  
AL LT LV MK RO SI

(72) Inventors:  
• Konstantinides, Konstantinos  
San Jose, California 95118 (US)  
• Tretter, Daniel R.  
No.18, Mountain View, California 94040 (US)

(30) Priority: 21.07.1998 US 120072

(74) Representative: Powell, Stephen David et al  
WILLIAMS, POWELL & ASSOCIATES  
4 St Paul's Churchyard  
London EC4M 8AY (GB)

(71) Applicant: Hewlett-Packard Company  
Palo Alto, California 94304-1112 (US)

### (54) ADAPTIVE VIDEO COMPRESSION USING VARIABLE QUANTIZATION

(57) An image compression system [50] for compound images containing both text and pictures is capable of receiving the images on a non-overlapping 8 by 8 pixel block and includes a discrete cosine transformer [14] connected to a quantizer [18] drawing lossy quantization factors from quantization tables [16]. The lossy quantization factors are modified by a variable quantization subsystem [54] based on the frequency of

changes in the block to provide low lossy quantization factors for high frequency of changes and high lossy quantization factors for low frequency of changes, the high frequency of changes being indicative of text and the low frequency of changes being indicative of pictures. The quantizer [18] is connected to an entropy coder using lossless entropy encoding factors from Huffman tables [22] to provide JPEG compliant files [26].

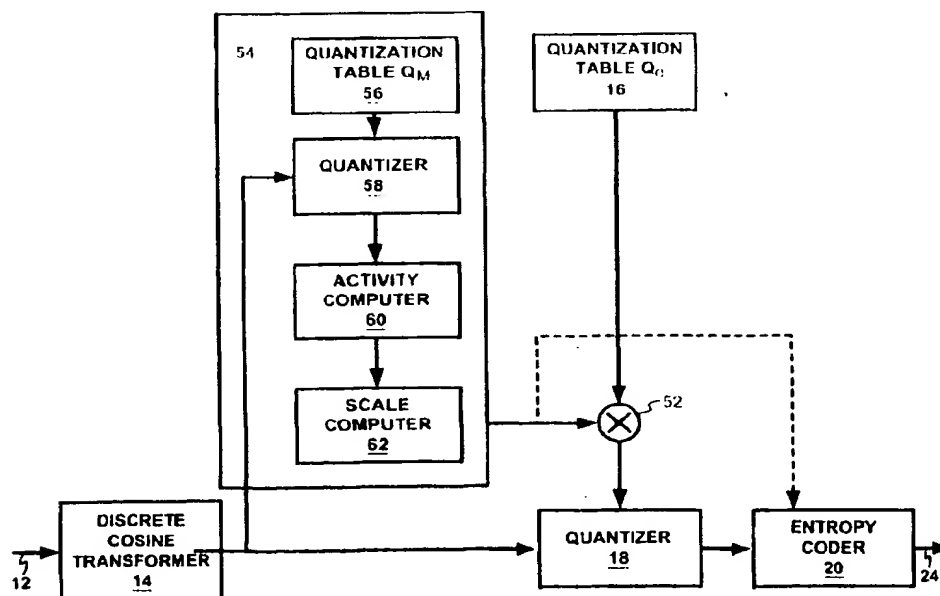


FIG. 3



European Patent  
Office

# EUROPEAN SEARCH REPORT

Application Number  
EP 99 30 4722

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
X	US 5 128 757 A (CITTA RICHARD W ET AL) 7 July 1992 (1992-07-07) * column 7, line 29 - line 54; figure 2 *	1-4,7	G06T9/00
X	SOON HIE TAN ET AL: "Classified perceptual coding with adaptive quantization" IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, AUG. 1996, IEEE, USA, vol. 6, no. 4, pages 375-388, XP002152024 ISSN: 1051-8215 * the whole document *	1-4,7	
A	BHASHARAN V ET AL: "Text and image sharpening of scanned images in the JPEG domain" PROCEEDINGS. INTERNATIONAL CONFERENCE ON IMAGE PROCESSING (CAT. NO.97CB36144), PROCEEDINGS OF INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, SANTA BARBARA, CA, USA, 26-29 OCT. 1997, pages 326-329 vol.2, XP002152025 1997, Los Alamitos, CA, USA, IEEE Comput. Soc, USA ISBN: 0-8186-8183-7		TECHNICAL FIELDS SEARCHED (Int.Cl.7) G06T
A	ROSENHOLTZ R ET AL: "Perceptual adaptive JPEG coding" PROCEEDINGS. INTERNATIONAL CONFERENCE ON IMAGE PROCESSING (CAT. NO.96CH35919), PROCEEDINGS OF 3RD IEEE INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, LAUSANNE, SWITZERLAND, 16-19 SEPT. 1996, pages 901-904 vol.1, XP000956371 1996, New York, NY, USA, IEEE, USA ISBN: 0-7803-3259-8		
The present search report has been drawn up for all claims			
Place of search THE HAGUE		Date of completion of the search 6 November 2000	Examiner Pierfederici, A
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document	

EPO FORM 1503 03/82 (P4C01)



**THIS PAGE BLANK (USPTO)**